Running head: FEMALE FACES AND BODIES

Key words: ugliness, beauty, prototypes, symmetry, skin condition, body color, decision making, averageness, stimulus complexity

Grammer, K., Fink, B., Juette, A., Ronzal, G. & Thornhill, R. (2001). Female faces and bodies: n-dimensional feature space and attractiveness. In: G. Rhodes & L.A. Zebrobwitz (Eds.). Advances in Visual Cognition. Volume I: Facial Attractiveness. Ablex Publishing. pp. 91-125.

Female faces and bodies:

N-dimensional feature space and attractiveness

Karl Grammer, Bernhard Fink, Astrid Juette, Gudrun Ronzal

Ludwig-Boltzmann-Institute for Urban Ethology

Althanstrasse 14

A-1090 Vienna

Austria

Randy Thornhill

Department of Biology

The University of New Mexico

Albuquerque, New Mexico 87131

USA

Summary

Many studies show that female attractiveness plays an important role in human mate selection. Research in the past has focused on the influence of single features, e.g., eye size or breast size, in attractiveness judgements. In recent years, bilateral symmetry and averageness or prototypic appearance have been discussed as possible general principles of attractiveness. The puzzle remaining is which features actually contribute to the perception of attractiveness and how are these integrated to result in attractiveness attribution. In this paper we propose that attractiveness perception and judgment takes place in a multidimensional feature space. If attractiveness signals mate quality honestly, the single features making up the multiple dimension should actually correlate positively and thereby compose a single ornament of mate value.

In a rating study, three sets of males (each n = 10) rated three views (face alone, nude back, and nude front with face covered) of digital images of women (n = 92) in Austria as well as in the USA. Symmetry, averageness, skin color, hair color, stimulus complexity and surface texture were assessed with digital image analysis. Thirty-six features on the digital images were measured by hand at anatomically defined points. A principal component analysis reveals that the n-dimensional feature space can be reduced to four main dimensions.

Computer simulations of the possible underlying cognitive decision making imply that a fast and frugal algorithm, which uses the rule "*simply avoid the worst*," best explains attractiveness ratings. Thus, beauty could be a negative concept, which finds its expression in the avoidance of ugliness.

Beauty Perception And Cultural Relativism

Charles Darwin wrote in 1872 in his book *The Descent of Man*: "It is certainly not true that there is in the mind of man any universal standard of beauty with respect to the human body" and "It is, however, possible that certain tastes may in the course of time become inherited, though I know of no evidence in favor of this belief."

Charles Darwin came to the conclusion that there are highly variable beauty standards in different cultures. He had asked missionaries and ethnographers to describe the beauty standards of different ethnic groups. The diversity of answers that he received made generality difficult. It still seems to many researchers that beauty standards are arbitrarily culturally determined. For example, Crogan (1999) points out that beauty standards vary highly over time and history and between (and even within) societies. This argument for a high degree of cultural relativism culminates in the following statement: "Evolutionary psychologists have failed to demonstrate convincingly that preferences for particular body shapes are biologically based. . . . Current data suggest that body satisfaction is largely determined by social factors, and is intimately tied to sexuality" (p. 164).

This argumentation, however, is contrasted with the fact that humans have a pre-occupation with attractiveness, and it has been repeatedly shown that we treat people differentially according to their physical appearance, preferring attractive others (Baugh & Parry, 1991; Zebrowitz, 1997; Hatfield & Sprecher, 1986).

Where does this obsessive pre-occupation with beauty and attractiveness come from? If some psychological algorithms were able to process information and solve problems better than others, and thus result in more offspring by Darwinian selection, then humans are quite likely to have widespread, even universal, adaptations in our thinking and reasoning (Cosmides, Tooby, & Barkow, 1992). This logic also applies to attractiveness ratings and might explain the human obsession with looks. In this paper we maintain that human attractiveness judgments have evolved by selection and are responsible for our perception of attractiveness and beauty.

Moreover, we argue that because beauty judgments are evolved, faces cannot be separated from bodies during judgments (Thornhill & Grammer, 1999). Attractiveness ratings seem to be an integral part of impression formation. When two people start to interact, this will typically take place from a distance. Thus, the perception of body shape might often be prior to the perception of faces. In this article we will argue that much research on facial attractiveness has the severe shortcoming of reducing the perception of attractiveness to a simplistic, one-dimensional concept.

In order to illustrate our view, we will first review different approaches to the measurement of beauty. We start with feature measurement approaches, then proceed to descriptions of averageness and symmetry, and also review the background of possible evolutionary explanation. We will do this in order to identify possible cues or signals, which could be used by raters in attractiveness judgments. Then we will try to link the beauty measurement approaches to a more general theory of decision-making involving the concept of attractiveness as a unidirectional, n-dimensional feature space in which multiple ornaments are organized signals. We treat women's beauty only.

Finally, we will empirically examine our approach in two different ways. One way is the consideration of single ornaments. The second approach involves examining naturally occurring groups of ornaments. We will test which of the approaches can produce the best prediction of actual attractiveness ratings.

What is physical in attractiveness?

Several rating studies, especially those by Iliffe (1960), have shown that people of the same ethnic group share common attractiveness standards. In these standards, ratings of pictures for beauty and sexual attractiveness seem to be the same over social class, age and sex. This work has been replicated several times by Henss (1987, 1988). Thus, beauty standards are shared in a given population.

Moreover, recent studies (Cunningham, Roberts, Wu, Barbee, & Druen, 1995) suggest that the constituents of beauty are neither arbitrary nor culturally bound. The consensus, on which a female is considered to be good-looking or not, is quite high in four cultures (Asian, Hispanic, and black and white women all rated by males from all cultures, see Cunningham, Rhodes, Dion, this volume). Three-month-old children gaze longer at attractive faces than at unattractive faces, which led Langlois, Roggman, & Reiser-Danner (1990) to conclude that beauty standards are not learned and that there is an innate beauty detector. If humans posses a common standard in attractiveness, there must be common features that are used for its detection and classification.

Single features approach

The first question is: What are the features of attractiveness and is it possible to find single features that are correlated with attractiveness? If attractiveness has a relation to mate-selection, then we would expect on theoretical grounds that the evaluation of traits in the opposite sex will include two categories of physical features: first, those traits should be valued that guarantee optimal reproduction and fertility (e.g., youth, Symons, 1979), and second, the features of certain sexual dimorphisms. If both sexes are adapted to optimal survival and reproduction, and if the constraints for both sexes are different, we will expect sexually dimorphic traits, which reflect each sex's adaptation. Thus, the signals that make a male typically 'male' and a female typically 'female' should be key signals for ratings of attractiveness in mate choice. Early approaches to measuring physical attractiveness were done by measuring various traits in faces and bodies, having these faces and bodies rated, and then comparing the traits to these ratings.

The almost automatic positive reaction to babyness proportions in faces (Fridlund & Loftis, 1990) has led to the assumption that babyness could be involved in attractiveness perception, because it could signal neoteny or youth and thereby gives rise to care-taking behavior. Jones (1996) shows that relatively neotenous female faces, i.e., faces that appear to be younger than the actual age of the face, are rated as more attractive by male raters from five populations. Attractiveness and youth ratings could rely on the influence (presence or absence) of sex hormones. Wildt and Sir-Peterman (1999) showed that age ratings of female pictures are different than actual age, and female attractiveness in their sample was estrogen dependent. Thin lips, high forehead and big eyes have been mentioned in many studies as traits of "babyness" (Rensch, 1963; Cunningham, 1986; Johnston & Franklin, 1993). It is felt that forehead height should be especially attractive because it is supposed to signal babyness and neoteny (Hess, Seltzer, & Shlien, 1965; Cunningham, 1986; Eibl-Eibesfeldt, 1997). Besides forehead height, brows situated high in the forehead and their curvature are supposed be an expressive signal, which signals a permanent "eye-brow-flash" and thus will give the face an attentive, open look (Cunningham, 1986; Eibl-Eibesfeldt, 1997). A small mouth and full lips are attractive, because they are supposed to demonstrate high estrogen levels and thus optimal fertility hormone profiles (Johnston & Franklin, 1993; Thornhill & Gangestad, 1993). Grammer and Atzwanger (1993), however, have shown that high cheekbones, as a sign of maturity, have to be added for a face to be viewed as attractive. These various traits appear to be the result of a high estrogen-to-testosterone ratio during puberty, with estrogen involved in capping the growth of bony structure in the face and body (Thornhill & Grammer, 1999).

The hormone argument can be applied to nose width and length, lower face length, chin length and jaw width, which retain their small size in women under relatively low levels of male sex hormones (e.g., testosterone) and relatively high levels of estrogens, giving rise to sex-typical hormone markers (Thornhill & Gangestad, 1993). High estrogen-to-testosterone levels also result in the enlargement of lips and the fat pad in the upper cheek area. These enlargements are analogous to estrogen-mediated fat deposition in breasts, thighs and buttocks.

Ellison (in press) has found that current estrogen levels in saliva correlate positively with conception probability across women when age is controlled. Thus, estrogen-based phenotypic effects like high cheekbones, a small lower face and the female typical body shape may honestly signal fertility.

Some single features of the body depend on body fat distribution, e.g., female breast and buttock size, and correlate with attractiveness for the other sex (Hess, Seltzer, & Shlien, 1965). In general, females have twice as much body fat as males and this fat is distributed differently on the body. Males have an android fat distribution in which fat is found preliminarily in the abdominal region. Females have a gynoid fat distribution with fat mainly found in the gluteofemoral region. In addition, female breasts consist of about 80 percent fat. Thus, fat distribution is highly sex-specific.

Firm breasts with small, lightly pigmented areolas, erect nipples, and the breast axis pointing upward and out in a V-angle are expected to be rated as attractive, because they are signs of young adulthood and therefore fertility (Symons, 1979). Compared to the adult male, the adult female body build is marked by relatively low values of shoulder width. Thus, the female-typical shoulder width is expected to be attractive, because shoulder girdle and associated musculature develop under the influence of testosterone.

On one hand, the distribution of body fat signals the ratio of estrogen to testosterone, since the dominance of one of these hormones is responsible for a typical female or male fat-distribution and body shape. On the other hand, the amount of fat in the female body is important for a stable level of female sex-steroids, and thereby affects female menstrual cycling (at least 25% of body weight is necessary for cycling, Frisch, 1975). Kirchengast and Huber (1999) showed that body fat distribution rather than the total amount of body fat correlates with the onset of pubertal hormonal activity as well as with the probability of successful conception in women participating in an artificial insemination program (see also Singh, 1993). Thus, sex-specific fat distribution (i.e., breast and buttock size) is directly related to reproductive success, and the amount of visible fat can predict whether a female is fertile or not, but overall weight is linked to fertility as well: heavy mothers have more children. Appreciation of heavier women in various cultures seems to depend on environmental stability (Anderson, Crawford, Nadean, & Lindberg, 1992). In unstable environments, body weight is linked to status and to attractiveness (Furnham & Baguma, 1994).

Traits that are developed under the influence of sex hormones may not only signal optimal hormone levels for reproduction. They may also be “honest signals” of parasite resistance and general mate quality. The handicap principle (Zahavi, 1975) explains the evolution of extravagant and costly structures in terms of honest signals, which can be afforded only by high-quality individuals. With regard to sex hormones, Folstad and Karter (1992) provide evidence that they affect negatively an individual’s immune system. It is likely that androgens and estrogens allocate energy and resources between reproductive functions and immunological defenses, giving rise to a tradeoff between reproduction and survival. To allocate energy into immunological defenses, infected individuals may be forced to reduce the costs of sexual displays by lowering their sex hormone levels (Wedekind and Folstad, 1994). Thus, traits that signal high sex hormone levels are honest in the sense that they impose an indirect handicap on the immune system.

Moreover, parasite resistance is a main theme in mate-selection (Hamilton & Zuk, 1982). The battle that hosts wage against parasites cannot be won by permanent adaptation. Whenever a host develops an effective immunity against parasites, the parasites will evolve countermeasures. Parasites have an edge because they have shorter life cycles than hosts. This theoretical consideration also would provide a more parsimonious explanation for the attractiveness of certain babyness features in women. Because these features develop under sex hormones, they may be signals in the context of the handicap principle instead of stimuli promoting care taking.

Another signal is the absence and presence of body hair, which is also sexually dimorphic and therefore likely a feature of attractiveness. Removal of female body hair is more common than removal of male body hair. Females appreciate body hair developed under male sex hormones, but males prefer its absence. Cleanly shaved legs and armpits are a youthful trait (Symons, 1979).

Males seem to prefer long head hair in women (Ronzal, 1996) and female hair growth on the head is more stable. Indeed, hair loss and baldness are a result of male sex hormones. Long hair thus is sexually dimorphic.

The general function of hair (on the head, in the armpits and pubic hair) may be the distribution of pheromones produced in the apocrine glands. Hair is expected to give a greater surface for pheromone distribution into the air. Thus, long female hair may be a “pheromone-distribution organ” correlated with optimal female sex hormone levels. The scalp (and all other regions of the human body that are covered with hair) has apocrine glands, which are thought to be responsible for pheromone production (Stoddard, 1995).

Female hair color is also a trait of attractiveness. Rich and Cash (1993) have shown that blonde hair, although infrequent in most populations, dominates in pictures of females presented in magazines for males. This gives rise to a completely different issue. Thelen (1983) showed that preference for any hair color (brunette, blonde or red) depends on the distribution of hair color in a population. Males prefer the “rare” color. This leads to an unstable situation and a given color cannot be consistently selected as an attractive trait. The reason for this situation might be a preference for “rare” genes, which could help in the host–parasite race discussed above.

Skin color is also a feature of attractiveness. Van den Berghe and Frost (1986) have proposed that change in skin color enables human males to distinguish fertile post-pubescent females from infertile pre-pubescent ones. Accordingly, men select lighter-skinned women, because they are presumably more fertile. One possible explanation of such selection is infantile mimicry: the imitation of infantile traits by one sex to reduce aggression and to elicit care-taking behavior in the other (Lorenz, 1943; Eibl-Eibesfeldt, 1997). This trend toward infantilization is more advanced in women than in men (Frost, 1988). Thus, skin traits may reliably indicate several aspects of female mate value.

The attractive average

There are, however, many objections to the approach of decoding attractiveness by using single features. One is methodological. Many researchers measure many features (up to a few hundred) and then correlate them with attractiveness. Oftentimes, there is no correction for the number of statistical tests done and hence the replication rate across studies is low. For many features, another objection against single-feature study for the assessment of attractiveness is that the relationship between size of some features and attractiveness ratings sometimes is curvilinear.

In the case of facial features, this has been shown using computer-generated averaged faces, an approach based on Galtons’ (1879) method of photographic averaging. Several studies have shown (Kalkofen, Müller, & Strack, 1990; Langlois, Roggman, & Reiser-Danner, 1990; Müller, 1993; Grammer & Thornhill, 1994; Rhodes & Tremewan, 1996) that computer-generated composite faces of women are more attractive than most of the single faces used for generating them[1].

Curvilinear relations also can be found among non-facial body features. An average waist-to-hip ratio of 0.70 is judged attractive (Singh, 1993). The range of waist-to-hip ratio (WHR) is similar for the sexes before puberty (0.85 to 0.95). After puberty, women's hip fat deposits cause their WHRs to become significantly lower than men's. Healthy, reproductively capable women have WHRs between 0.67 and 0.80, whereas healthy men have a ratio in the range of 0.85 to 0.95.

The preference for averageness could be explained in terms of stabilizing selection where selection is against the extremes of a population (Symons, 1979; Langlois and Roggmann, 1990). Thornhill and Gangestad (1993) proposed that preference for average trait values could have evolved, because on continuously distributed heritable traits, the average connotes genetic heterozygosity. Heterozygosity could signal an outbred mate or provide genetic diversity in defense against parasites. Parasite resistance should be a valued signal in mate choice (Hamilton & Zuk, 1982).

Langlois and Roggman (1990) emphasized the importance of prototypes in human categorization and recognition as the basis for the attractiveness of averageness. Indeed, a basic feature of human cognition is the creation of "prototypes" (Rosch, 1978). This means that we constantly evaluate stimuli from our social and non-social environment and classify these stimuli into categories. This reduces the amount of information, increasing its usefulness and its economy of storage in memory. It is still unclear whether average faces are prototypes—that is, are average representations of stimuli of one class. Composites look familiar precisely because they are average or representative faces, resembling many faces. If our brain uses prototypes, the averageness of composites made from each sex might well correspond with being "prototypical" male or female.

In our view, one of the most interesting possibilities from attractiveness prototyping is that we may be able to adjust adaptively our beauty standards to the mean of the population.

If average faces correspond to prototypes, attractive averageness expands our possibilities of mate-choice. If we had an inflexible biological template ("Suchbild") for attractiveness judgment, we could run the danger of either never meeting somebody fitting the template or being frustrated by the non-fitting mates we find. Through prototyping, our beauty standard could be adjusted to the population we live in. In view of this, we should expect learning mechanisms to be involved in beauty standards, and an ability to adjust these standards to the population we live in. This could also explain the apparently different standards of aspects of beauty over time and cultures (Grammer, 1993).

There are single faces, however, that are not average, but are rated very attractive (Grammer & Thornhill, 1994; Perrett, May, & Yoshikawa, 1994). These findings have provoked considerable debate, suggesting alternative explanations for the attractiveness of composites. The higher attractiveness ratings of computer-generated composites over many individual faces could be due to the fact that the prototypes lack skin blemishes and are much softer than normal pictures. Thus, higher attractiveness ratings for the composites could be caused by their apparent skin condition.

There is a relationship between dermatoses and elevated levels of sex hormones, which is often correlated with an ovarian dysfunction (e.g., "polycistic ovary syndrome," Schiavone, Rietschel, & Sgoutas, 1983; Steinberger, Rodriguez-Rigau, Smith, & Held, 1981). To our knowledge, no one has reported a study that directly measured skin texture or experimentally changed skin texture by blurring or removing artifacts in order to examine its effects on attractiveness. Furthermore, the effects of image manipulation on the perception by raters are unknown, because the effects of the manipulations themselves on the surface of the face are not measured directly, and most researchers assume that the manipulations affect all stimuli in the same way. Fink & Grammer (under review) focused on these issues. They used female faces that were standardized for their shape and found that skin texture significantly influences attractiveness ratings. Langlois, Roggman, & Musselman (1994), however, argued against the blurring artifact by showing that a composite of different photographs of the same face is not particularly attractive. They also felt that composites look more familiar than individual component faces, but saw this as a part of their averageness explanation (Langlois et al., 1994).

In our view, the debate on the effect of averageness on attractiveness ratings is still unresolved. Still another problem is that composite faces are more symmetric than most of the individual faces used for creating them.

The role of symmetry

Bodily and facial asymmetries manifested in normally bilaterally symmetric traits are the result of developmental instabilities during embryonic and subsequent development (Thornhill & Gangestad, 1996). Developmental instability arises during development as the result of environmental perturbations such as pathogens and of genetic perturbations such as deleterious mutations. As a consequence, symmetry may be an honest or reliable signal of an individual's genetic and phenotypic quality and thereby important in mate choice (Thornhill & Gangestad, 1993). Indeed, symmetry as a mate-selection criterion has been shown in many species, from scorpionflies (Thornhill, 1992) to birds (Møller & Pomianowski, 1993) and humans.

Grammer and Thornhill (1994) provided the first evidence that facial symmetry may be a significant feature in facial attractiveness. These findings, which have been referred to as the 'symmetry hypothesis' (Rhodes, Sumich, & Byatt, 1999), have been disputed. Langlois et al. (1994) were critical of the hypothesis after finding that symmetry did not correlate with attractiveness, and that perfectly symmetric faces were less attractive than the original ones. Other studies (Kowner, 1996; Samuels et al., 1994; Swaddle and Cuthill, 1995) also showed that asymmetric faces were judged more attractive than symmetric faces. Most of these studies have created stimulus faces by aligning one half of a face with its mirror reflection. Perrett et al. (1999) showed that this mirror reflection could introduce abnormal feature shapes. Swaddle and Cuthill (1995), as well as Rhodes et al. (1999), produced their stimuli by combining a face with its mirror image. Perrett et al. (1999) criticized this method because it increases the amount of blemishes in the skin texture. They used a warping technique involving calculation of symmetrical landmarks. In a rating study, these faces were compared to the original asymmetric faces. They found that the more symmetrical faces were rated more attractive. Thornhill and Gangestad (1999a) have critically reviewed the research on facial symmetry and attractiveness and conclude that symmetry itself plays a significant role in attractiveness judgments. In addition, there is some evidence that bodily symmetry correlated with ratings of facial attractiveness (see Thornhill & Gangestad, 1999a).

Sensory exploitation

Preferences for symmetrical stimuli also could be subsumed under the concept of "*sensory exploitation*." Recent research suggests that symmetry is one of the main factors in recognition of, and reaction to, stimuli. Several computer simulations have shown that neural networks, when confronted with various stimuli, respond better and easier to symmetry (Enquist & Arak, 1994; Johnstone, 1994; Enquist, this volume). Symmetry in faces may lead to an easier encoding of faces. Also, symmetric faces may be easier to process because of their lower complexity. Thus, symmetrical stimuli may exploit the sensory system of the receiver and therefore be attractive in contexts unrelated to mate choice. Such general preferences for symmetry (Rensch, 1963) would serve no obvious mate-selection function.

Similarly, average or prototypical faces might have a fit better to prototypical neural templates. As a result, prototypes might be recognized faster, easier and more completely, and thus might create higher nervous excitation. This could be the reason for a preference towards average stimuli. Maybe our brain accepts better fitting stimuli more willingly. Müller (1993) has called this kind of process "*neuroaesthetics*."

Yet it is still unclear how the features involved are combined and interact during attractiveness judgments. Considerable evidence has accumulated in recent years supporting the hypothesis that both facial and bodily physical attractiveness are health certifications and thus represent honest signals of phenotypic and genetic quality (Thornhill & Grammer, 1999; Thornhill & Gangestad, 1999). The hypothesis that beauty connotes health was first proposed by Westermarck (1921) and later by Ellis (1926) as well as Symons (1990, 1995). There is no doubt, regardless of how variables are decoded, that attractiveness is a cognitive construct in the eye (brain, mind) of the beholder.

Beauty: an n-dimensional feature space?

On theoretical grounds, the many features in the female face and body that affect attractiveness can be reduced to three categories. The first is signals of optimal sex-hormone levels and/or immunocompetence handicap. The second is signals of developmental stability. These categories may be linked via parasite resistance. The third is youth and thus fertility.

A salient question is how are the features that influence attractiveness related to each other? These features are analogous to the ornaments of birds in that they are involved in sexual attraction. Møller and Pomianowski (1993) identify three hypotheses to explain multiple sexual ornaments among species. One is the *multiple message hypothesis*, which argues that each ornament signals a specific, unique property of the condition of an individual. This could be the case if one ornament signals resistance to disease X, another to disease Y, and neither disease or only one affects overall condition and hence lifetime reproductive success. The second is the *redundant signal hypothesis* that multiple ornaments more reliably predict overall condition. In this case, mate choosers should pay attention to several sexual ornaments because, in combination, they provide a better estimate of general condition than does any single ornament. The third is the *unreliable signal hypothesis* that most ornaments of species with multiple sex traits do not signal condition.

Cunningham et al. (1995, 1997) suggested a 'multiple fitness model'—that attractiveness varies across multiple dimensions, rather than a single dimension, with each feature signaling a different aspect of mate-value. It would correspond with the first hypothesis above.

Thornhill and Grammer (1999) showed that independent attractiveness ratings in Austria and the U.S.A. of the same women in each of three poses (face, front nude with faces covered, and back nude) are significantly positively correlated, as predicted by the redundant signal hypothesis. Because of the connection of the attractiveness features of the face, back and front to estrogen, the correlation between the ratings of the different pictures implies that women's faces and bodies comprise what amounts to a single ornament of honest mate value. Also, the redundant signal hypothesis implies that features involved in different communicative channels should relate to overall condition. In order to study this, Rikowski and Grammer (1999) focused on the question of whether body scent signals general mate quality like other cues in sexual attraction. The results showed significant positive correlations between facial attractiveness and sexiness of body odor for female subjects (also see Thornhill & Gangestad, 1999b). Moreover, the more symmetric the body of a woman, the sexier is her smell (but see Thornhill & Gangestad, 1999b).

Schleidt and Crawley (1980) provided an approach that might solve the problem of integration of many features or signals into one meaning. They proposed an 'n-dimensional vector approach' to multiple signal communication, where each signal is a single vector with size and direction.

If such an n-dimensional feature space is present in the perception and attribution of attractiveness, its structure can be assessed to test hypotheses about multiple ornaments. If the redundant signal hypothesis is true, all features will point in the same direction and the receiver will be able to decode the signal of overall condition clearly and unmistakably by assessing multiple ornaments. The n-dimensional feature space would then be uni-directional. This will not be the case for the multiple message hypothesis. If the redundant signal hypothesis is true, a possible method for decoding attractiveness itself, given that the ornaments have the same signaling quality, would be to compare their magnitude. This is the simplest approach we can imagine, because it does not require any further processing like shape analysis or the detection of the relationship of one feature to another feature. This method would not work when all ornaments signal different qualities of condition, as in the multiple message hypothesis.

Gigerenzer and Goldstein (1996) suggest that people use fast and frugal algorithms, which produce the same results as more complex designs of decision-making algorithms in many every-day decision-making problems. One example of a simple algorithm in attractiveness judgments would be "*the worst (or best) feature approach*." This means that signal receivers simply compare the size of the best or worst feature in an n-dimensional feature space (regardless of the feature's content) in order to come to a decision that one person is more attractive than another. Note that this method only takes the size of the feature into account, not its quality. Such fast and frugal algorithms would work only if the redundant signaling hypothesis is true. This would allow for subtle attractiveness discriminations, because the sizes of single features will differ in any n-dimensional space. Such an algorithm would suggest that beauty perception may not be a positive concept, but the reverse concept: "*avoid ugliness*."

Yet even more simple methods are conceivable. When there is no direct comparison available, a simple threshold model could be used. Then the worst feature has to be over a certain threshold before the whole person is rated attractive. This method, too, would allow for fine-tuned discriminations according to the size of the feature that exceeds the threshold.

Basically, all cognitive models are constrained in two ways. The first is complexity, which is related to utility. The principle of parsimony would suggest that simpler models have a higher explanatory value than more complex models. Second, any model has to use knowledge in order to come to a decision. As with the first constraint, we can assume that the less knowledge a model uses, the higher is its utility.

In order to explore the possible existence of a uni-directional, n-dimensional feature space for the assessment of female beauty, we examine the following:

We evaluate the contributions of the physical features discussed above to attractiveness ratings of women. We will use traditional measuring techniques with landmarks and newly developed image analysis techniques.

We test how these features are organized either as multiple, different signals or as redundant signals. The multiple message hypothesis states that different ornaments signal different properties of the condition of an individual. The redundant signal hypothesis predicts that in multiply ornamented species all traits used in mate choice will show condition-dependent expression and will be correlated with overall condition (Møller & Pomianowski, 1993).

We also examine these hypotheses by comparing the variability of the measured traits to the attractiveness ratings. A relation of low variability of the traits to high attractiveness would support the redundant signaling hypothesis.

Furthermore, we conduct a principal component analysis on the attractiveness traits. The resulting factors provide insight into the organization of the sexual traits in the n-dimensional feature space.

Finally, we try to build a model of how attractiveness ratings are reached by the brain, based on the assumption that simple fast and frugal algorithms are used for beauty ratings. This is done by simulating possible cognitive strategies and then examining their fit to the attractiveness ratings. This is a new approach to research in decision making for attractiveness judgments. The basic assumption is that decoding is easier if there is a uni-directional, n-dimensional feature space with consistent stimuli (i.e., if one trait is attractive, all traits are attractive). It is exploratory and does not provide a final solution to the question of how attractiveness ratings map to decision-making in attractiveness judgments. It provides direction as to how future research might approach the questions of this decision-making.

Methods

Subjects and rating procedure

Akira Gomi took nude photographs of 92 Caucasian women, ranging in age from 18 to 30. The women responded to Gomi's advertisement in the *Los Angeles Times*, were paid about $50 (U.S.), and signed a consent form allowing their photographs to be used commercially or in scientific studies. The photos were taken under constant light conditions with a digital camera and there was no color processing involved. Gomi standardized the photos for size and distance to the camera. Facial photos were with neutral expression and faces appeared to have little to no make-up on them. Body photos were with standardized posture (standing upright, arms extending down the sides of the body with the feet a few inches apart) and perpendicular orientation to the camera.

Three different picture poses of each subject were rated for attractiveness: face only, front of nude body from head to knee with head and hair blocked out, and nude back from head to knee bend. The photos (350 x 480 pixels size, 72 dpi resolution, color control by

Color Sync from Apple Computers) were presented to each rater on a 17" computer screen using a presentation software developed by the first author. The pictures were first presented individually for five seconds and sequentially to each rater in order to give the rater an overview of the photos. Immediately after this preview, the photos were presented to the rater for rating. Each rater conducted ratings privately without anyone else in the room. Photo order in presentations was randomized initially, but all raters of a set of photos saw the same order in both the preview and actual rating aspects. Rating was on a 1–7 scale of attractiveness, where 7 is most attractive and 1 is least. Opposite sex attractiveness ratings of facial photographs are known to be related positively to romantic and sexual interest in the person depicted (e.g., Grammer, 1993).

Each of the three sets of pictures of each subject were rated by men who self-reported their age (mean age = 25 years, range = 19–55) and ethnicity. Each rater rated only one of the three sets. Three groups of men (N = 10 per group) rated each set in each country; thus, there were 30 raters in Vienna and 30 raters in New Mexico. The New Mexico ratings were used to examine the cross-cultural generality of the attractiveness ratings. All Viennese raters identified themselves as Caucasian, but U.S. raters showed a mixture of self-reported ethnic backgrounds: Oriental (3 raters), Hispanic (5 raters), American Indian (3 raters), and Caucasian (19 raters). (For inter-rater reliability, see Results.)

In some pictures the visibility of focal body parts was restricted by hair. These photos were not usable for measurements and digital image analysis and therefore the final number of women used for this paper (n = 70) differs from that used by Thornhill and Grammer (1999).

Measurements

Standardization

Before image-processing operations were performed, the pictures were standardized to the same orientation. The faces were coded with 51 landmarks (source co-ordinates, see Figure 1). In a second step, the mean co-ordinates (destination co-ordinates) for all faces and the mean respective landmarks were calculated. A computer program for the size standardization of the faces and figures was developed following Bookstein (1997) and Wolberg (1990) using a simplified version of the procrustes approach. This program calculates the center of gravity of the source co-ordinates for each face. The face was then moved on the picture plane so that the center of gravity of the face fell on the center of gravity of the destination co-ordinates. Finally, each face was resized to 150% of its original size. Faces were then scaled down in one-pixel steps, until the square sum of the difference between source and destination co-ordinates reached a minimum (least square method). After scaling, the face was rotated about the center of gravity for 45°. Then the same method as above was applied for stepwise rotation. This resulted in non-distorted size- and orientation-optimized pictures in relation to the center of gravity of the face. The same procedure was applied to the front (46 landmarks) and the back views (28 landmarks) (Figure 1). Landmark placement reliability was tested with an untrained student and resulted in a mean error of 1.5 (S.D. = 0.8) pixels placement error. This is an error of 0.43% picture size in the horizontal direction and 0.31% in the vertical direction.

Feature measurements

Thirty-six features were measured. Some basic measurements (height, weight, breast size, and waist and hip circumference) were taken directly by Akira Gomi. All other measurements were taken from the pictures themselves.

The features were all derived from the literature as having a potential influence on attractiveness ratings in faces or bodies. Thus, for each feature, we could predict the possible relationship to attractiveness (see Introduction and Appendix). The feature measurements were divided in three main categories: (a) direct pixel dimensions of features (see Appendix, number 1 to 10), (b) measurements analyzed automatically by digital image analysis features (see Appendix, number 11 to 16) and (c) global stimulus descriptors derived from digital image analysis features (see Appendix, number 17 to 18).

Direct pixel dimension measurement was used for those features in which the size of the feature is predicted to correlate either negatively or positively with attractiveness. This involved 19 features (forehead height, brow height and curvature, lower face length, nose width and length, chin length and jaw width, eye width and height, mouth width and height, cheek bone height, breast height, areolar size, angle of breast axis, buttock size, hair length and shoulder width).

Digital image analysis methods were applied to the measurement of 11 additional features. These features are face, breast and buttock asymmetry; skin texture homogeneity; skin color for face, front and back; hair color and pubic hair color; nipple color; and lip color. For all eleven measurements predictions were derived from the literature.

Finally we added six global stimulus descriptors to the measurements (averageness for face, front and back; stimulus complexity for face, front and back). The stimulus complexity features allow us to assess whether there is a global influence on the ratings of the stimulus itself as predicted by neuroaesthetic theory. The digital assessment of averageness allows us to avoid the pitfall of digital alterations of pictures in order to assess the influence of averageness (see endnote 1).

Results

Trait measurements and attractiveness

Thornhill and Grammer (1999) showed using the same stimulus materials and attractiveness ratings as in this study that there are significant correlations for the ratings of face with back and with front poses, as well as for the poses back and front. Austrian and American males agreed in their attractiveness ratings, both within and between countries. We calculated the reliability for the subsample (n = 70) used herein. The Austrian ratings for face, front and back have a reliability of Cronbach's alpha = 0.71. The American ratings showed a comparably high reliability of alpha = 0.72. When all ratings are analyzed together (Americans, Austrians, faces, front and back), alpha was 0.87. This indicates that the attractiveness ratings for all views and the two cultures are quite uniform. In the subsample, the raters judged the attractiveness of a face similarly to that of the front and the back view of the same person (Spearman correlations between the three views: face and front $r = 0.29$, $p < .05$; face and back $r = 0.37$, $p < .05$; front and back $r = 0.74$, $p < .05$) This suggests a uni-directional, n-dimensional feature space, as predicted by the redundant signal hypothesis. Thus, attractiveness was summed up for all views (total attractiveness, Table 1).

In order to test the contribution of the single measurements to the overall rating of attractiveness, we calculated the single correlations between each of the variables outlined above and total attractiveness (see Table 1); 12 of 36 correlations reached significance and 19 pointed in the predicted direction. Three of the significant correlations (breast circumference, back view color value and front view color value) pointed in the opposite direction than predicted. Given the high number of correlations, their statistical significance is not the most salient issue. Instead, the size of a correlation coefficient has to be taken into account. Also, even when there is a weak relation between the size of the single measurements and attractiveness, it appears that facial and body traits, in general, contribute to the ratings in the same direction.

Insert Table 1 about here

Next, we tested for the uniformity of the features. For this, the data were z-transformed in order to make the values for the different traits comparable in size. For subsequent interpretation of the results, this transformation means that we do not look any more at the absolute size of a trait for one person compared to all the sizes of that trait for others in the sample. We correlated the variability of values of all traits with total attractiveness. The result was a small, but significant correlation, $r = -0.23$ ($\underline{n} = 70$, $\underline{p} = .05$). This means that a female who was rated attractive showed low variability in those traits we selected for the analysis. In a female who was rated unattractive, all traits tended to differ more in value and extremes may occur. This pattern is consistent with the existence of a uni-directional, n-dimensional feature space, where attractive features in the bodies and faces are linked within women, and extreme combinations of traits within women are rated unattractive. We turn now to the question of what heuristics a person might use for the assessment of attractiveness. There are several strategies possible and we simulate these strategies with our data.

Single feature measurements and decision-making

As discussed above, there is some evidence for the redundant signal hypothesis. Thus, we explored whether single traits could be the basis of decision making in attractiveness ratings. The first strategy we tested is "*avoid the most extreme trait*." This is a simple decision strategy, which would work as follows. The rater would look at all traits, compare them and then take the most extreme one (for instance, nose size) for rating and comparison between rated subjects. This strategy would not require a uni-directional, n-dimensional feature space, but it would require knowledge about the variation of all traits in order to decide which trait is of extreme size in comparison to all other traits. In order to test this, the maxims of the absolute values of the z-scores for each target were calculated, and the value of the trait with the highest absolute z-score was taken as the most extreme trait for the simulation. The results in Table 2 indicate that such a strategy would lead to poor decision accuracy. Although there was a tendency for the sizes of the trait with the highest z-score to be negatively correlated with attractiveness, none of the correlations were significant. Thus, attractiveness ratings do not appear to be attributed by an avoidance of extremes strategy.

The second strategy would be as simple as the strategy above. This strategy can be called an "*idiosyncratic choice*." In this strategy, only one trait is looked at, a haphazard or random choice, but the rater has knowledge of the attractiveness distribution for this trait. In this case, a rater uses, for instance, only eye size, and another rater uses only mouth width. To examine this, we multiplied all features' sizes that correlate negatively with attractiveness by -1, so that all correlations with attractiveness are positive. Then we selected for each rater one variable at random. Each rater would rely on just one of the features and use it to make his attractiveness decision. This was done in 15 simulations. The correlation between randomly selected trait size and attractiveness reached significance once out of 15 simulations. Table 2 shows the results as a correlation range. This strategy also performed poorly in predicting attractiveness.

Insert Table 2 about here

Another strategy can be called "*use the best trait*." In contrast to the first strategy and like the second strategy, a uni-directional, n-dimensional feature space has to be present. This means that all traits are redundant signals with regard to attractiveness. In order to simulate this, all trait values correlating negatively with attractiveness were multiplied by -1. The strategy then would work as follows: compare all traits and then select the best one (i.e., trait with maximum cue value, the trait correlating highest with attractiveness), and use the cue value of this trait for the attractiveness decision. As shown in Table 2, the correlation between actual ratings and trait size by this method is higher than with the first strategy but is not significant.

"*Use the best*" represents a positive concept of attractiveness. In contrast, the forth strategy we tested is an avoidance strategy. In this case, attractiveness ratings are perceived as negative concepts, not positive ones. Beauty judgment might not be an appraisal of positive stimuli, but an avoidance strategy, here called "*avoid the worst*." It would use the same knowledge as the third strategy, but this time simply takes the minimum size of the trait with the smallest value for a decision. In Table 2 we show that, with the exception of the face, such a strategy yields significant correlations between trait selection results and attractiveness ratings. Thus, so far, the avoidance strategy provides the best fit to data.

One problem with the single feature approach is that it might not be feasible because there also could be curvilinear relations between certain features and attractiveness. Consequently, we suggest that an approach that subsumes single features would be more appropriate for decision making about attractiveness in a heuristic way. For this reason, we tried to identify one or more latent variables behind the single features by factor analysis. First, however, we will look at two other alternatives to a single feature approach: averageness and stimulus complexity.

The role of averageness

One objection against a single-features strategy is the amount of information that has to be used for decision-making. Thus, we tried to simulate information reduction strategies of prototyping and determining stimulus complexity. The prototyping strategy is called "*compare to averageness*." In order to achieve this, each picture was morphed to the mean co-ordinates and then the amount of deviance from the mean was calculated. When a picture is warped to other co-ordinates, each pixel will move by a certain distance. These distances were summed up. The measure describes how far away a picture is from the mean of all pictures, i.e., from the assumed prototype. Figure 2 shows morphed prototypes for the face, front, and back views.

Insert Figure 2 about here

In addition to examining the relation between prototypicality of a stimulus and attractiveness, we determined whether there is a correlation between face, front and back view deviance from the average. If so, this would be evidence for a uni-directional, n-dimensional feature space. There was indeed significant positive correlations between deviance of face and front (n = 70, rs = .25, p = .04), deviance of front and back (n = 70, rs = .29. p = .01), but there was no significant correlation between face and back (n = 70, rs = .06, n.s.). This result proves our assumptions only partially.

When we compared the three views with the respective ratings, we found negative correlations for deviance of face and facial attractiveness (n = 70, rs = -.26, p = .05), for deformation of front view and front attractiveness (n = 70, rs = -.36, p = .002), but again none for back view and back attractiveness ($\underline{n}$ = 70, $\underline{rs}$ = -.02, n.s.). Finally, we computed a complete deviance and correlated it with attractiveness, but only front attractiveness deviance from average reached significance (see Table 2).

These results suggest that an average appearance might be attractive, but there is probably more than a simple fitting on an average template. They also imply that attractiveness is more than just averageness.

Stimulus complexity and attractiveness

An assumption sometimes made about attractiveness ratings is that the mere complexity of a stimulus could be responsible for the ratings. Thus, we tested another strategy: "*use the least complex stimulus*," and calculated the amount of image compression, which is possible for each picture. The more compression possible, the less complex the stimulus. As was the case for deviance from prototype, there were significant positive interrelations for compression of face, front and back view ($\underline{n}$ =70, face-back $\underline{rs}$ = .25, $\underline{p}$ = .03; face- front $\underline{rs}$ = .11, n.s.; back - front $\underline{rs}$ = .54, $\underline{p}$ = .001). Moreover, there was a positive correlation between variability (see above) and total complexity ($\underline{n}$ = 70, $\underline{rs}$ = .41, $\underline{p}$ < .001). When we looked for the correlations of the single complexity measures and attractiveness, there was only one significant correlation: the less complex the stimulus, the higher the total attractiveness (see Table 2).

Factor analysis of the traits and attractiveness

So far, the analysis shows that one strategy (i.e., "*avoid the worst*") accounts for the highest variance reduction in attractiveness ratings. The disadvantage of this strategy is that it requires a large amount of knowledge; the rater has to know each trait, its distribution and its relation to attractiveness (i.e., either positive or negative).

Our final approach was to apply an information reduction procedure in order to explore the organization of the feature space. A principal components factor analysis (varimax rotation) was carried out on all 36 measurements. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy equals 0.59. This is mediocre, but not unacceptable. A four-factor solution explains only 56% of the variance, but factor loadings are reasonably high (see Table 3). This result might reflect that the choice of variables is exploratory.

Insert Table 3 about here

In order to visualize the factors, the 10 pictures with the highest regression scores on one factor were morphed together. Factors 3 and 4 are easy to interpret. Factor 4 (row d in Figure 3) reflects nubility. Smaller eyes and a lower forehead contrast with high values for breasts, breast axis, mouth width, and height. Also note the high loadings of face and breast symmetry. Factor 3 (row c in figure 4) seems to be a weight and size factor. Values are high in weight, breast, hip and waist measures. This factor could be interpreted as a general body mass factor. These signal combinations on the senders' side have a correspondence in the attractiveness ratings on the receivers' side. Factor 4 correlates significantly positively with attractiveness and Factor 3 correlates significantly negatively (Table 3).

Insert Figure 3 about here

Factors 1 (Babyness) and 2 (Color) are mainly dominated by color values. Factor 1 loads high on color values, and moderate on asymmetrical breasts, eye size, small buttocks and a small nose. This factor correlates significantly negatively with total attractiveness (see Table 3). Factor 2 contains light colors and a symmetrical face. This factor does not correlate with attractiveness (Table 3). These results indicate that the feature space has an internal structure were different features are linked and which have a positive or negative relation to attractiveness. This would greatly simplify processing processing**.**

When we apply the best of or worst of strategy to the four regressed factor scores with Factor 1 and Factor 3 reverted, we find relatively high correlations for the positive decision (use the best factor) and the negative decision (use the worst factor, see Table 2). This suggests that there are at least two templates—an attractiveness template and an unattractiveness template—used for decision-making.

As a final step, we wanted to know whether there is any connection between the deviance from the mean and these templates. This is not the case—correlations range from 0.04 to 0.13. Only complexity shows a significant correlation to the regressed factor scores from Factor 1 ($r_s = .25$, $p = .04$). This result indicates that it is possible to extract trait combinations from the sender, which are rated as either attractive or unattractive.

Discussion

The results of our approach show that it is possible to extract features that are related to meta-theoretical evolutionary explanations for trait values in attractiveness ratings. Not all traits correlate in the predicted ways, but many do. However, the absence of some predicted correlations does not prove that there is no general relationship in the predicted direction.

We show that the n-dimensional feature space of attractiveness is coherent and uni-directional, i.e., features tend to point in the same direction, and low variability in attractiveness between features of a woman is attractive. This is difficult to explain with cultural relativistic theories of attractiveness. If it is an arbitrary learning paradigm, why then should the sender signal coherently?

When we look at the single correlations of traits with attractiveness, we find several cases where the direction of the correlation contradicts correlations found in other studies. One trait where the actual correlation contradicts the predicted direction is breast circumference. If we take a closer look at the data, this measure might be confounded by breast form because breast form with large distances between lower breast fold and nipple correlates positively with attractiveness. Another result we did not find was correlations of attractiveness with all measures that are supposed to be a part of the babyness scheme. This replicates the findings of Grammer and Atzwanger (1994) on a different data set, where they assumed that babyness features are associated with incompetence and childishness. This is why males possibly would avoid those features and prefer traits of maturity.

Color values were analyzed in an attractiveness study for the first time. The results do not support our prediction. Lighter skin is not preferred. The males in our study preferred darker, apparently possibly tanned skin in this sample. Tanning of skin is also associated with more homogenous skin. There could be an interaction between these two parameters.

Our results indicate that bodies and faces can be considered as a whole. The fact that a uni-directional, n-dimensional feature is suggested allows us to use all traits (and their values) from face and body in simple fast and frugal algorithms in order to come to an attractiveness decision. But this decision-making approach may coexist with other strategies. The results indicate that simple strategies like avoidance of extremes, idiosyncratic learning or simple deviations from the average explain the least variance in decision making. In order to come to a decision, it seems to be necessary to have knowledge about the relation between traits and attractiveness. The question, then, is if and how this knowledge is gathered. In our approach, this could be solved easily, as long as the sender signals coherently in a uni-directional, n-dimensional feature space. If this n-dimensional feature space does not exist, it would be much harder to gather or accumulate any knowledge on what is attractive. This means that the starting point for any learning and flexible adjustment of attractiveness templates to an existing population (this has to be the case in order to maintain the chances of finding a suitable partner) could be provided by the coherently signaling sender. The sender's appearance would foster learning in an "*attractiveness direction*."

The only approach without the need of some knowledge prerequisites that worked was the use of stimulus complexity. This complexity could be related to variability and thus non-coherence in the n-dimensional feature space. This is one of the most interesting results.

We did not find a relation between total averageness and attractiveness, although we were able to replicate the finding by Grammer and Thornhill (1994) that facial averageness correlates positively with attractiveness.

Another major result is that it seems to be possible (although the factor analysis is preliminary) to create attractiveness templates from body and facial features, i.e., the sender's whole appearance is coherent. We do not assume that the prototypes we generated with morphing are the actual templates, nor that prototypes like this are represented in the brain; they are simple examples in order to visualize the findings. This result is not the foundation for a new constitution typology. It is simply the construction of prototypes to show that our approach is possible.

Basically, we show that taking knowledge on all traits into account would work as reliably as using prototypes. But, in both cases, we find a surprising result. Current literature describes beauty as a positive concept, but a negative concept works as well, if not better. If this is the case, research should consider redefining beauty as the "*avoidance of ugliness*." When we take the principle of parsimony into account, we could favor the prototype solution.

This solution would require less knowledge than a single trait solution where the relation between all traits and attractiveness has to be known to the decision maker.

Our study itself has considerable shortcomings. Primarily, we conducted a simple correlation analysis. No corrections were applied for the number of tests carried out. The variance of the sample might not be high enough in some traits, and there might be curvilinear relations between traits and attractiveness ratings. We also did not use proportional measures. This was done for methodological reasons in order to keep the amount of measured distances small and comparable in size. Actually, the prototype approach would not create problems with curvilinear relations and proportions, because a deviation or fitting on the prototype is measured. In fact, there also could be other strategies for information reduction, which use parallel distributed processing. There are still other alternatives like an hierarchical decoding approach where features are processed one after the other, as, for instance, take the best, then take the worst, and so on. We will address these possible strategies using different stimulus material in our forthcoming work. Future directions of research might benefit from use of 3D-meshes or wire frames, instead of predefined distances, in order to avoid possible bias in distance selection.

References

Anderson, S. & Anger, S. (1995). *PC-Graphik fuer Insider*. München: SAMS.

Anderson, J. L., Crawford, C. B., Nadean, J., & Lindberg, T. (1992). Was the Duchess of Windsor right? A cross-cultural review of the sociobiology of ideals of female body shape. *Ethology and Sociobiology, 13*, 197-227.

Bässmann, H. & Besslich, P. W. (1993). *Bildverarbeitung Ad Oculus*. (2nd Ed.) Berlin: Springer Verlag.

Baugh, S. G., Parry, L. E. (1991). The relationship between physical attractiveness and grade point average among college women. *Journal of Social Behavior and Personality, 6*, 29-228.

Bookstein, F. L. (1997). *Morphometric Tools for Landmark Data; Geometry and Biology*. Cambridge, UK: Cambridge University Press.

Bruce, V. (1988). *Recognizing Faces*. London: Erlbaum.

Collins, M, & Zebrowitz, L. (1995). The contributions of appearance to occupational outcomes on civilian and military settings. *Journal of Comparative Psychology, 25*, 129-163.

Corballis, M. C. (1991). *The Lopsided Ape: The Evolution of the Generative Mind*. New York: Oxford University Press.

Cosmides, L., Tooby, J., & Barkow, J. H. (1992). Introduction: Evolutionary psychology and conceptual integration. In: J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind* (pp. 3-18). New York: Oxford University Press.

Crogan, S. (1999). *Body image. Understanding Body Dissatisfaction in Men, Women and Children*. London: Routledge.

Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi experiments on the sociobiology of female beauty. *Journal of Personality and Social Psychology, 50*, 925-935.

Cunningham, M. R., Roberts, A. R., Wu, C.–H., Barbee, A. P., & Druen, P. B. (1995). Their ideas of beauty are, on the whole, the same as ours: Consistency and variability in the cross-cultural perception of female attractiveness. *Journal of Personality and Social Psychology, 68,* 261-279.

Darwin, C. (1872). *The Descent of Man and Selection in Relation to Sex*. London: John Murray.Davis, L. S., Johns, A. & Aggarwal, J. K. (1979). Texture analysis using generalized co-occurrence matrices. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1, 3*, 251-259.

Dhawan, A. P. & Sicsu, A. (1992). Segmentation of images of skin lesions using colour and texture information of surface pigmentation. *Computerized Medical Imaging and Graphics, 16,* 163-177.

Eibl-Eibesfeldt, I. (1997). *Die Biologie des menschliche Verhaltens*. Weyarn: Seehamer Verlag.

Ellis, H. (1926). *Studies in the Psychology of Sex*, Vol. IV. Philadelphia: F.A. Davis.

Ellison, P. T. Reproductive ecology and reproductive cancers. In: C. Panter–Brick, & C. Worthman (Eds.), (in press). *Hormones, Health and Behavior*. Cambridge: Cambridge University Press.

Farah, M. J. (1990). *Visual Agnosia: Disorders of Object Recognition and What They Ttell Us About Normal Vision*. Cambridge, MA: MIT Press.

Fink, B. & Grammer, K. (under review). Skin texture and female facial beauty.

Fisher, R. A. (1930). *The Gentical Theory of Natural Selection*. London: Oxford University Press.

Friedlund, A. J., & Loftis, J. M. (1990). *Sex-Differences in Smiling to Babies and Animals: Do Females Show a Greater Preference for Juvenescence?* Unpublished manu-script.

Frisch, R. E. (1975). Critical weights, a critical body composition, menarche and the maintenance of menstrual cycles. In: E. S. Watts (Ed.), *Biosocial Interrelations in Population Adaption*. The Hague: Mouton.

Frost, P. (1988). Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine, 32*, 1, 38-58.

Furnham, A. & Baguma, P. (1994). Cross-Cultural Differences in the Evaluation of Male and Female Body Shapes. *International Journal of Eating Disorders, 15 (1),* 81-89.

Galton, F. (1879). Composite portraits, made by combining those of many different persons in a single resultant figure. *Journal of the Anthropological Institute, 8*, 132-144.

Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Grammer, K. (1995). *Signale der Liebe: Die Biologischen Gesetze der Partnerschaft.* München: Deutscher Taschenbuch Verlag.

Grammer, K & Atzwanger, K. (1994). Der Lolita-Komplex: Sexuelle Attraktivität und Kindchenschema. In: K.-F. Wessel & F. Naumann (Hrsg.), *Kommunikation und Human-ontogenese: Berliner Studien zur Wissenschaftsphilosophie und Humanontogenetik*. Band 6. Bielefeld: Kleine Verlag.

Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology, 108*, 233-242.

Haralick, R. M., Shanmugam, K. & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3, 6*, 610-621.

Hatfield, E., & Sprecher, S. (1986). *Mirror, Mirror . . . : The Importance of Looks in Everyday Life.* Albany: State University of New York Press.

Henss, R. (1987). Zur Beurteilerübereinstimmung bei der Einschätzung der physischen Attraktivität junger und alter Menschen. *Zeitschrift für Sozialpsychologie, 18*, 118-130.

Henss, R. (1988). „…wer ist der / die Schönste im ganzen Land?“ Zur Beurteilerübereinstimmung bei der Einschätzung der physischen Attraktivität. *Annales – Forschungsmagazin der Universität des Saarlandes, 1*, 54-58.

Hess, E. H., Seltzer, A. L., & Shlien, J. M. (1965). Pupil response of hetero- and homosexual males to pictures of man and women. *Journal of Abnormal Psychology, 70*, 165-168.

Iliffe, A., H. (1960). A study of preferences in femine beauty. *British Journal of Psychology, 51*, 267-273.

Johnston, V.S. & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology*, *14*, 183-199.

Jones, D. (1996). *Physical Aattractiveness and the Theory of Sexual Selection*. Ann Arbor, MI: Museum of Anthropology, University of Michigan.

Johnstone, R. A. (1994). Female preferences for symmetrical males as a by-product of selection for mate recognition. *Nature, 372*, 172-175.

Kalkofen, H., Müller, A., & Strack, M. (1990). Kant´s facial aesthetics and Galton´s composite portraiture—are prototypes more beautiful? In: L. Halasz (Ed.), *Proceedings of the 11th International Congress on Empirical Aesthetics* (pp.151-154). Budapest: International Association for Empirical Aesthetics.

Kenrick, D. T. & Gutierres, S. E. (1980). Contrast effects and judgements of physical attractiveness: When beauty becomes a social problem. *Journal of Personality and Social Psychology, 38*, 131-140.

Kenrick, D. T., Gutierres, S. E., & Goldberg, L. (1989). Influence of erotica on ratings of strangers and mates. *Journal of Experimental Social Psychology, 25*, 159-167.

Kirchengast, S. & Huber, J. (1999). Body composition characteristics, sex hormone levels and circadian gonadotropin fluctuations in infertile young women. *Collegium Anthropologicum, 23(2),* 407-423.

Langlois, J. H., Roggman, L. A., & Reiser-Danner, L. A. (1990). Infant's differential social responses to attractive and unattractive faces. *Developmental Psychology, 26*, 153-159.

Locher, P., Unger, R., Sociedade, P., & Wahl, J. (1993). At first glance: Accessibility of the physical attractiveness stereotype. *Sex Roles, 28, 11/12*, 729-743.

Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung. *Zeitschrift für Tierpsychologie, 5,* 235-409.

Møller, A. P. & Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology, 32*, 167-176.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience, 9(5),* 555-604.

Müller, A. (1993). Visuelle Prototypen und die physikalischen Dimensionen von Attraktivität . In: R. Niketta, M. Hassebrauck (Eds.), *Physische Attraktivität*. Göttingen: Hogrefe.

Perrett, D. I., May, K. A. & Yoshikawa, S. (1994). Facial shape and judgement of female attractiveness. *Nature, 386,* 239-242.

Rensch, B. (1963). Versuche über menschliche Auslösermerkmale beider Geschlechter. *Zeitschrift für Morphologische Anthropologie, 53,* 139-164.

Rhodes, G., Brennan, S, & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representation of faces. *Cognitive Psychology, 19,* 473-497.

Rich, M. K., & Cash, T. (1993). The American image of beauty: Media representations of hair color for four decades. *Sex Roles, 29 (1/2),* 113-123.

Rikowski, A., & Grammer, K. (1999). Human body odour, symmetry and attractiveness. *Proceedings of the Royal Society of London, Series B, 266*, 869-874.

Ronzal, G. I. (1996). *Physische Charakteristika weiblicher Schönheit*. MA thesis Natural Sciences, University of Vienna, Austria.

Rosch, E. H. (1978). Principles of categorization. In: E. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-47). Hillsdale, NJ: Erlbaum.

Schiavone, F. E., Rietschel, R. L., Sgoutas, D., & Harris, R. (1983). Elevated free testosterone levels in women with acne. *Archives of Dermatology, 119,* 799-802.

Schleidt, W. M., & Crawley, J. N. (1980). Patterns in the behaviour of organisms, *Journal of Social and Biological Structure, 3*, 1-15.

Singh, D. (1993). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of Personality and Social Psychology, 65*, 293-307.

Steinberger, E., Rodriguez-Rigau, L. J., Smith, K. D., & Held, B. (1981). The menstrual cycle and plasma testosterone level in women with acne. *Journal of the American Academy of Dermatology, 4(1),* 54-58.

Symons, D.(1995). Beauty is in the Adaptations of the Beholder: The Evolutionary Psychology of Human Female Sexual Attractiveness. In: P.R. Abramson and S.D. Pinker (Eds.), *Sexual Nature/Sexual Culture* (pp. 80-118). Chicago: University of Chicago Press.

Symons, D. (1979). *The Evolution of Human Sexuality.* Oxford: Oxford University Press.

Thelen, T. H. (1983). Minority type human mate preference. *Social Biology, 30,* 162-180.

Thornhill, R. (1992). Fluctuating asymmetry and the mating system of the Japanese scorpionfly (*Panorpa japonica*). *Animal Behaviour, 44*, 867-879.

Thornhill, R. & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human Nature, 4*, 237-269.

Thornhill, R. & Gangestad, S. W. (1996). The evolution of human sexuality. *Trends in Ecology and Evolution, 11*, 98-102.

Thornhill, R. & Grammer, K. (1999). The body and face of woman: One ornament that signals quality? *Evolution and Human Behavior, 20*, 105-120.

Thornhill, R. & Gangestad, S.W. (1999). The scent of symmetry: A human sex pheromone that signals fitness? *Evolution and Human Behavior, 20*, 175-201.

Valentine, T. (1991). A unified account of the effects of ditinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43(A)*, 161-204.

Van den Berghe, P. L. & Frost, P. (1986). Skin color preference, sexual dimorphism and sexual selection: a case of gene-culture co-evolution? *Ethnic and Racial Studies, 9*, 87-118.

Westermarck, E. (1921). *The History of Human Marriage*. London: MacMillan.

Wildt, L. & Sir-Peterman, T. (1999). Oestrogen and age estimations of perimenopausal women. *Lancet, 354*, 224.

Wolberg, G. (1990). *Digital Image Warping*. Los Alamitos, California: IEEE Computer Society Press.

Zahavi, A. (1975). Mate selection. A selection for a handicap. *Journal of Theoretical Biology, 53*, 205-214.

Zebrowitz, L.A. (1998). *Reading faces. Window to the soul?* Boulder, Colorado: Westview Press.

Appendix

The notation will be as follows: if distances are used, the points are presented; xP1xP3 would indicate width of right eye for example (See Figure 1a). If regions are used, the notation is top/left/bottom/right; 27/2/3/28 would describe the rectangle where hair color has been measured (see Figure 1c). A value behind a point number describes a positive or negative offset to this point in pixels. For example 27(10)/2(-10)/3(-10)/28(-10) would indicate that the rectangle from above was scaled down by a value of 10 pixels. Zero point for the co-ordinate system is the upper left corner of the picture.

Features: pixel dimension from photographs.

1) Forehead height (A: ((P21y-P14y)+(P23y-P14y))/2).
2) Brow height (Figure 1a: P21y-P16y) and curvature (Figure1a: ((P3x-P16x)+(P19x-4x))/2).
3) Lower face length (Figure1a: P13y-P3y), nose width (Figure1a: P8x-P7x) and length (Figure1a: P29y- ((P3y+P4y)/2)), chin length (Figure1a: P13y-P29y) and jaw width (Figure1a: P10x-P9x
4) Eyes (width (Figure1a: ((P3x-P1x)+(P2x-P4x))/2) and height (Figure1a: ((P22y-P21y)+(P24y-P23y))/2)).
5) Mouth width (Figure1a: P12x-P11x) and height (A: P35y-P32y), cheekbone height (Figure1a: $(((P13x\text{-}P5x)^2+(P13y\text{-}P5y)^2)^{0.5} + ((P6x\text{-}P13x)^2+(P13y\text{-}P6y)^2)^{0.5})/2)$.
6) Breast height (Figure1b: ((P11y-P9y)+(P12y-P10y))/2. High values would indicate high amount of body fat and thus optimal sex hormone profiles (Frisch, 1975).
7) Areola size (Figure1b: ((P51y-P49y)+(P52y-P50y))/2) and angle of breast axis (Figure1b: P11x-P9x)+(P10x-P12x)).
8) Buttock size (Figure1c: P15y-P12y).

9) Hair length (Figure1c: P26y-P1y).

10) Shoulder width (Figure1b: P4y-P3y).

Features: Digital image analysis

11) Face, breast and buttock asymmetry. (Face: Figure1a: P16/P1/P35/P2; Breast: Figure 1b: P7/P7/P12+10/P8; Buttock: Figure 1c: P12/P13/P25+10/P14)) High values are unattractive (see introduction). This method was done with a digital image analysis algorithm for the detection of asymmetry. The program developed by the first author created a window on the facial picture that was defined by left and right outer eye corner, top of the brows, and the lower lip. This window then was divided in _n_ horizontal 1 pixel wide symmetrical slices. Each slice was then moved to the left for 50 pixels and moved back in one-pixel steps. For every step the difference between left and right part of the slice was calculated. The symmetry point is reached when the difference between the two halves of each slice reaches a minimum. Thus _n_ symmetry points were calculated as the minimum difference between the sum of pixels of the left and the right half of the respective slice. In an ideal symmetrical face the line through all symmetry points is a straight line and equals the distance between top of the brows and lower lip. The symmetry index was calculated as the length of the symmetry line divided by the height of the window. This results in measures of asymmetry greater 1, where 1 depicts absolute symmetry. Compared to the distance measure this method does not only measure the symmetry of tissue and bone structure as expressed in landmarks - it also uses skin disturbances for determination of asymmetry, which might indeed play a crucial role for the perception of symmetry.

12) Skin texture. (Face: Figure1a: P22/P1/P35/P2) The region was the center of the face. Co-occurrence-matrix: For the analysis of texture, classification and image segmentation, we

used a method based on co-occurrence-matrices (for the source code see Bässmann & Besslich, 1993 (pp. 150-155), Haralick 1973; Davis, Johns & Aggarwal, 1979; Dhawan & Sicsu, 1992). Co-occurrence-matrices are based on the spatial relationships of pairs of grey values of pixels in digital texture images. They count, how often pairs of grey levels of pixels, which are separated by a certain distance and lie along a certain direction, occur in a digital image of texture. Co-occurrence-matrices are not used directly but features based upon them are being computed. The aim is that these features capture some characteristics of textures, such as homogeneity, coarseness, periodicity and others. Out of Haralicks' original 14 features we chose homogeneity. High values of homogeneity are supposed to be attractive. Texture was measured in a region of interest in the central face.

13) Skin color (Face: Figure1a. P22+20/P1/P7/P3; Front: Figure1b: P14+15/P14/P14/P15; Back: Figure1c: P10+15/P10/P10/P11). We determined skin color in an RGB- and separately in an HSV-color space. Color was measured by calculating the mean hue and value. Colors in the RGB color space are represented as values between 0 and 65535 for red, green and blue. The hue of the HSV color space component is an angular measurement, analogous to a position in a color wheel. A hue value of 0° indicates the color red; the color green is a value corresponding to 120° and the color blue is at a value corresponding to 240°. The value component describes brightness or luminance. A value of 0 represents black; a maximum value means that the color appears brightest. For more information on the use of HSV color spaces see 'Advanced color imaging on the MacOS' (Apple Computer Inc., 1995). Body brightness, back brightness and front brightness was measured by calculating, value of the following picture regions: Face color in a rectangle on the right cheek determined by left and right corner of the eye, the point of maximum cheekbone protrusion and the end of the nose. Back and front in a rectangle 15-pixel wide

reaching from left to right waist in the front view and in the back view. High values are supposed to be attractive.

14) Hair color (value) (Figure1c: P27/P27/P27+20/P28), pubic hair color (value) (Figure1b: P20-10/P20-10/P20/P20+10).

15) Nipple color (Figure1b: P9-5/P9-5/P9+5/P9+5). Light, red areolars and nipples are supposed to signal nubility (Symons, 1979). High values of the red component and brightness are supposed to be attractive.

16) Lip color (Figure1a: P34/P34-10/P35/P34+10). High values of the red component and brightness are supposed to be attractive.

Global stimulus descriptors (face only)

17) Averageness (all views) (Face: Figure1a: P16/P15/P13/P20; Front: Figure1b: P7/P7/P16/P17; Back: Figure1c: P6/P6/P25/P7). Amount of necessary form warp to obtain the average picture. The mean co-ordinates for all females were calculated. Then each face was form warped to these co-ordinates using an algorithm by Anderson and Anger (1995). If a face is average, then the individual pixels do not move very much, but if the form deviates considerable from average, many pixels have to be moved in order to reach the average form. Necessary pixel movement was calculated as the sum of all movements in pixels performed to reach average form. When averageness is considered attractive, low values would indicate attractiveness (see introduction).

18) Stimulus complexity (all views) (Face: Figure1a: P16/P15/P13/P20; Front: Figure1b: P7/P7/P16/P17; Back: Figure1c: P6/P6/P25/P7). The complexity of a stimulus is represented in the amount of possible compression of the picture. The higher the stimulus complexity, the less it can be compressed. Here we used Run-length encoding. This is a simple compression approach, which seeks strings of like pixels within the image and

assigns a single value. High complexity would result in larger file sizes. High values are predicted to be attractive because they would indicate low complexity and low complexity is more easily processed (see introduction).

Table 1

Pearson correlations of single features with assessed attractiveness of the face, the front view, the back view and all together. For prediction of the direction of the correlation (+/-) see introduction. Measurements, which were taken from the real person are marked with †, otherwise measurements were taken from photographs (N = 70).

	Attractiveness				
	Face	Front view	Back view	Total	Prediction
Height†	.02	.17	.31*	.24*	+
Breast circumference†	-.21	-.18	-.23	-.24*	+
Waist circumference†	-.05	-.31*	-.46*	-.36*	-
Hip circumference†	-.01	-.23	-.24*	-.20	+
Weight†	.01	-.14	-.18	-.14	-
Forehead height	-.18	-.14	-.19	-.22	+
Brow height	.07	.26*	.23	.23	+
Brow curvature	.13	.09	.02	.08	+
Lower face length	-.09	-.03	-.01	-.04	-
Nose length	-.10	.01	-.01	-.03	-
Nose width	-.13	-.02	.10	.02	-
Jaw width	.11	-.31*	-.18	-.15	-
Eye width	.25*	-.07	.05	.09	+
Eye height	-.29*	-.19	-.23	-.28*	+
Mouth width	.01	.10	.15	.12	-
Mouth height	.33*	.35*	.37*	.41*	+
Cheekbone height	-.08	-.06	-.20	-.12	+
Breast size	.06	.47*	.33*	.35*	+
Areolar size	-.16	-.45*	-.42*	-.41*	-
Angle of breast axis	.07	.10	-.01	.05	+
Buttock size	-.07	-.06	-.25*	-.18	+
Hair length	.18	.26*	.33*	.34*	+
Shoulder width	-.05	-.15	-.21	-.18	-
Face symmetry	.22	.25*	.32*	.34*	+
Breast symmetry	.14	.11	.18	.16	+

Buttock symmetry	-.09	.05	.03	-.01	+
Skin homogeneity	.16	.11	.03	.12	+
Face color value	-.02	-.09	-.09	-.10	+
Back view color value	-.19	-.26*	-.29*	-.31*	+
Front view color value	-.15	-.28*	-.20	-.25*	+
Hair color value	-.03	-.01	.00	-.01	+
Pubic hair color value	.21	.24*	.21	.26*	+
Areolar red color value	.18	.03	.09	.11	+
Areolar color value	-.19	-.03	-.11	-.11	+
Lip red color value	-.04	.00	-.03	-.02	+
Lip color value	-.07	-.14	-.10	-.13	+

*$p < 0.05$

Table 2

Correlations between outcomes of decision strategies and attractiveness ratings (N = 70).

Strategy	Attractiveness ratings			
	Face	Front	Back	Total
Avoid the most extreme trait	-.12	-.19	-.07	-.14
Idiosyncratic choice	-.24 -> .13	-23 -> .29	-.16 -> .36	-.20 -> .31
Use the best trait	.11	.19	.25*	.22
Avoid the worst trait	.20	.44*	.37*	.41*
Compare to average	-.13	-.28*	-.16	-.21
Use the least complex stimulus	-.15	-.19	-.21	-.24*
Use the best prototype	.13	.44*	.55*	.48*
Use the worst prototype	.35*	.49*	.45*	.53*

*p < 0.05

Table 3

The four factors, their Eigenvalues, the loading of the single traits on the factors and the rank-spearman correlations of the attractiveness ratings with the factors (N= 70).

	Babyness	Color	Obesity	Nubility
Eigenvalue	5.19	3.77	3.05	2.65
Areolar color value	.88	-.20	.11	.17
Areolar red color value	-.88	.20	-.12	-.18
Front view color value	.85	-.17	.03	.02
Back view color value	.84	-.19	.05	.04
Breast symmetry	-.68	-.10	.10	.28
Brow height	-.32	-.04	-.06	.01
Nose width	-.26	.09	-.17	.15
Lip color value	.16	-.90	-.02	.15
Face color value	.20	-.89	-.01	.15
Lip red color value	-.11	.89	.06	-.22
Skin homogeneity	-.06	.73	.06	.16
Face symmetry	-.02	.50	.02	.43
Brow curvature	-.08	.38	-.06	.05
Shoulder width	.09	.13	.10	-.06
Weight	.07	.01	.84	.14
Hip circumference	.06	-.02	.78	-.12
Waist circumference	.18	.08	.68	-.08
Breast circumference	.16	.03	.60	-.04
Buttock size	-.27	-.13	.57	-.23
Cheekbone height	.26	.22	.43	.16
Areolar size	-.14	-.14	.40	-.11
Jaw width	-.04	.23	.35	.20
Buttock symmetry	-.20	.26	-.35	-.18
Hair length	-.23	-.17	-.29	.06

Eye width	.19	.25	-.27	-.07
Pubic hair color value	-.01	-.06	-.19	.67
Breast size	.21	-.03	-.00	.55
Hair color value	-.05	-.12	.05	.53
Body height	.18	-.03	.36	.47
Angle of breast axis	-.21	-.11	.10	.43
Lower face length	.10	.27	.21	.37
Mouth width	.04	.20	-.18	.35
Eye height	.35	.17	-.10	-.35
Nose length	.04	-.00	.07	-.31
Mouth height	-.00	.19	-.17	.24
Forehead height	.13	.01	.01	-.14
	Correlations of factors with attractiveness ratings			
Face	-.17	.14	-.09	.23
Front	.18	.06	-.30*	.39*
Back	-.23	.05	-.43*	.37*
Total	-.24*	.10	-.37*	.41*

*p < 0.05

Figure Captions

Figures 1 a-c. This figure shows the landmarks on the three views (face, front and back) used in this study.

Figure 2 a-c. This is the average morph of all 70 females from this study. Pictures were morphed in one single pass for all 70 subjects.

Figure 3 a-d. This shows the morphs of the 10 females in each factor that loaded highest in the regression scores of the respective factor.

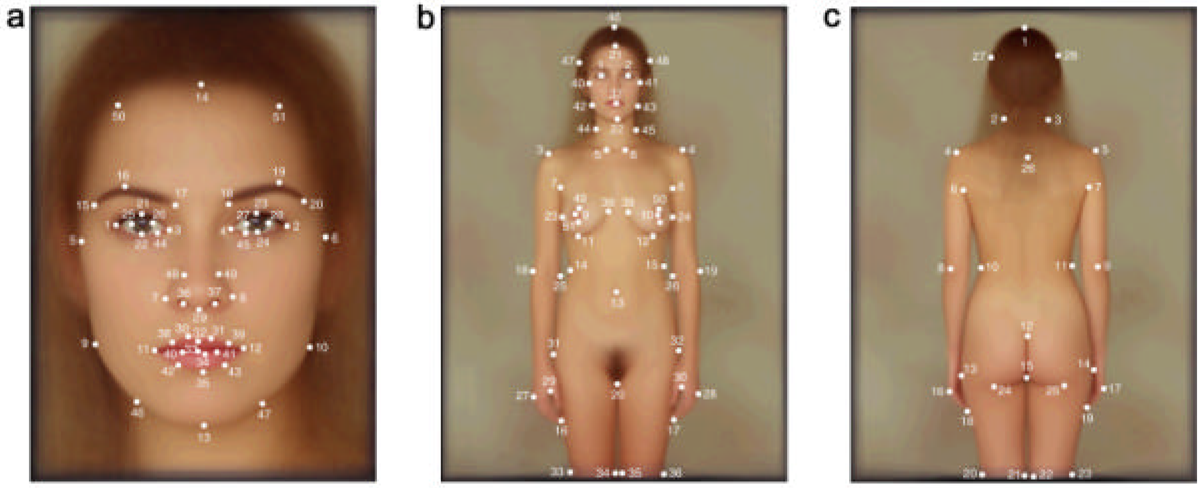

Figure 1

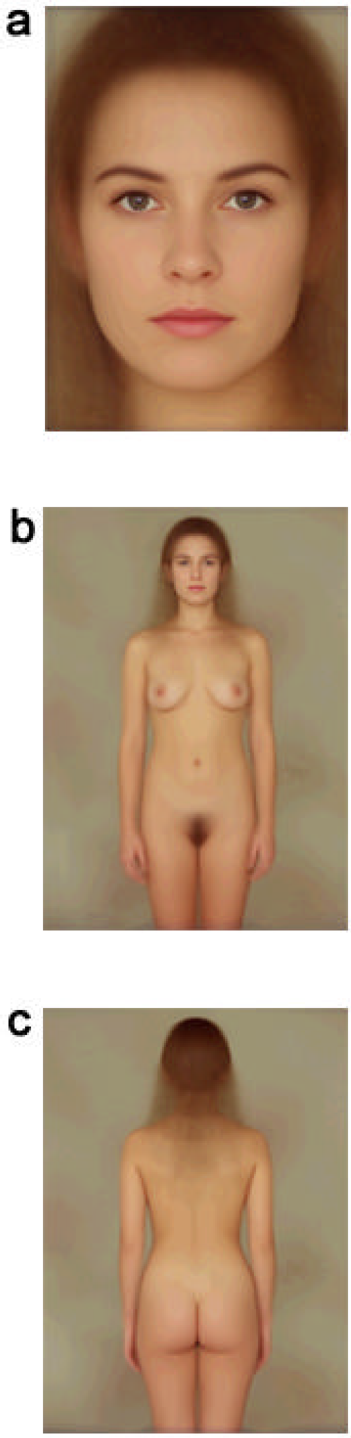

Figure 2

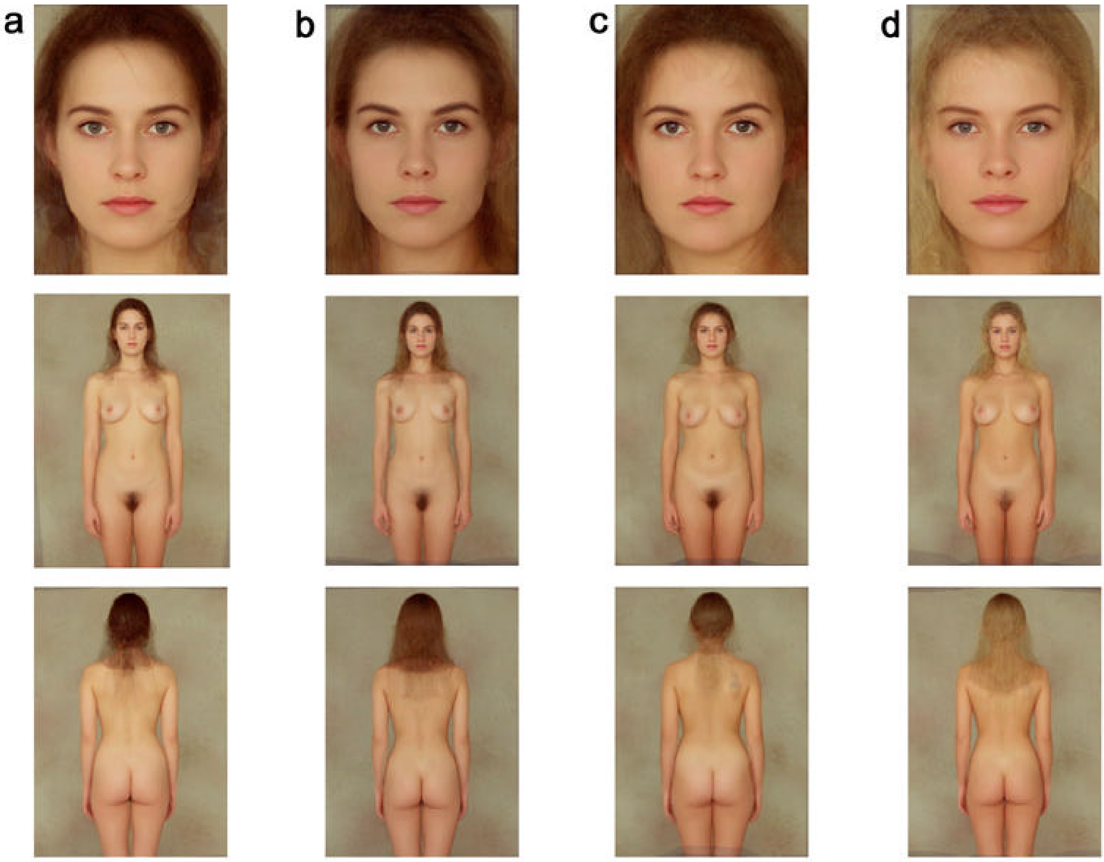

Figure 3

Endnote

[1] There are a lot of methodological problems with the making of computer composites. Thus, some prior research is questionable. The first problem is scaling of the faces using non-standard methods which result in distortion, although there are morphometric procedures which have been tested and verified (see Methods section). The second problem is in the method of morphing itself. Exact morphing requires establishing many landmarks. The reliability for placement of landmark point locations often is not reported in published studies. We would assume that the placement of landmarks may sometimes be biased, because it is done by hand, and often not by a person who is blind to the hypothesis. Third, when commercial programs are used, the exact algorithms used for warping and morphing are often unknown and undocumented. Therefore, it may be completely unclear whether artifacts were created. Fourth, calculating the mean value of two colors (even on a grayscale) requires transformation of the colors in a color space. This can be done in many different ways, and in many programs it is unclear how this conversion is achieved. We must assume that color correction is done in many cases, leaving it unclear if the actual result is the mean or not. Fifth, the nature of the applied morphing algorithms leads inevitably to an artifact, which could be responsible for the high attractiveness of composites. Usually, there are differences between neighboring pixel values in a picture. By calculating the mean values, these differences disappear and are smoothed out. Thus, a morphed picture is reduced in terms of three dimensionality in that it is flatter than the original faces. This is an artifact, which has to be controlled for by raising the picture contrast. Sixth, this applies to color and grayscale pictures equally. The composite will automatically have less contrast and be lighter than the individual pictures. None of the studies reported in the literature addresses this problem.